כדי לתמוך בשתי סוגי הקבצים גם CSV וגם Parquet בניתי פונקציית עזר בשם load\_data שבודקת את סיומת הקובץ וטוענת אותו בהתאם. בצורה הזו שאר הקוד יוכל תמיד לעבוד בלי להתחשב בסיומת הקובץ.

יש מספר יתרונות של parquet על csv כמו:

* מהיר יותר לקריאה וכתיבה
* גודל הקובץ קטן יותר
* שומר תמיד על סוגי הנתונים
* מותאם לעבודה עם Big data

קבצי csv אמנם פשוטים יותר לשימוש אך הרבה פחות מועילים...